DOCUMENT RESUME

ED 337 959 EC 300 701

AUTHOR DeStefano, Lizanne
TITLE Evaluating Effectiveness: A Comparison of Federal Expectations and Local Capabilities for Evaluation among Federally-Funded Model Demonstration Programs.
PUB DATE 91
NOTE 22p.
PUB TYPE Reports - Evaluative/Feasibility (142)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Accountability; Demonstration Programs; *Disabilities; Evaluation Needs; *Evaluation Problems; *Federal Programs; *Federal Regulation; Outcomes of Education; Program Effectiveness; *Program Evaluation; Secondary Education; Transitional Programs
IDENTIFIERS *Office of Special Educ Rehabilitative Services

ABSTRACT

This paper examines the issues arising when a set of federally funded model demonstration projects, the secondary curricula/student transition projects of the Office of Special Education and Rehabilitative Services, were obligated by federal initiative to improve the amount, quality, and use of evaluation. These issues included: (1) the nature of federal expectations and needs for evaluation data, including emphasis on outcomes and quantitative data in evaluation, the overpromising of program goals at project initiation, and funding cycles; and (2) the extent of local capabilities to meet those expectations and provide evaluative data, including problems of lack of evaluative expertise and objectivity at the local level and the uncertainty of rewards associated with project response to federal evaluative needs. Regulatory and methodological implications of the discrepancy between federal expectations and local capabilities along with implications for evaluation technical assistance are discussed. (13 references) (JDD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

X This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

ED337959

Evaluating Effectiveness: A Comparison of Federal Expectations and Local Capabilities for Evaluation Among Federally-Funded Model Demonstration Programs

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Lizanne DeStefano

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Lizanne DeStefano
University of Illinois at Urbana-Champaign
210 Education Building
1310 S. Sixth St.
Champaign, IL 61820
(217) 333-2245

EC 300701

BEST COPY AVAILABLE

Abstract

The scenario of a federal funding mandating the conduct and reporting of program evaluation activities has become increasingly more prevalent in education over the last two decades. This paper examines the issues that arose when a particular set of federally-funded model demonstration projects, OSERS' secondary/transition projects, were obligated by federal initiative to improve the amount, quality, and use of evaluation. These issues are organized into two areas: 1) the nature of federal expectations and needs for evaluation data; and 2) the extent of local capabilities to meet those expectations and provide evaluative data. Regulatory and methodological implications of the discrepancy between federal expectations and local capabilities along with implications for evaluation technical assistance are discussed.

Evaluating Effectiveness: A Comparison of Federal Expectations and Local Capabilities for Evaluation Among Federally-Funded Model Demonstration Programs

Perspective

It is commonly accepted that persons at different positions within or surrounding an organization have different values, purposes, and levels of understanding regarding an evaluation of that organization. The presence of these differing perspectives raises questions relating to their impact on evaluation such as: What effects do these differing perspectives have on the nature and quality of evaluating activity? Can an evaluation that is conducted to serve the purposes of decision makers at one position adequately address the needs of those at another? When there is an extraorganizational catalyst for the evaluation effort, such as a funding agency, what constraints are placed on the project's truthfulness? When the funding agency is the federal government, what perceived and actual expectations are held for the type and quality of evaluation data?

The scenario of a federal funding source mandating the conduct and reporting of program evaluation activities has become increasingly more prevalent in education over the last two decades. In recent years, the period of federally commissioned external evaluations of large scale federally funded programs appears to have given way to the practice of national aggregation of findings of locally conducted individual evaluation studies.

Since the passage of The Education for All Handicapped Children Act (EHA, P.L. 94-142) in the mid-seventies, administrators of special education have become accustomed to providing evaluation data to federal and state agencies. The prescriptive nature of the law and its accompanying regulations have resulted in the widespread conduct of program evaluation as a monitoring activity, measuring the extent to which individual programs are in compliance with state and federal regulations concerning appropriate identification, placement in the least restrictive environment, individualized program planning, due process procedures, and equal access. The types of evaluation data traditionally reported

include: head counts of pupils served (by handicapping condition), documentation of existence of required services, and statements detailing the procedures used to access them.

Over time, this emphasis on compliance or accountability monitoring in special education has come to be misinterpreted as a measure of program quality and effectiveness (DeStefano, 1990). This misinterpretation is so prevalent from both local and federal perspectives, that a compliant program is usually synonymous with an effective one. In reality, the processes associated with evaluating compliance and effectiveness are quite different. The main purpose of compliance monitoring is to ascertain the extent to which existing programs are meeting legislative requirements in terms of populations served and service delivery. Compliance monitoring does not address how effective a program is at achieving its goals nor does it include attempts to assess systematically the direct or indirect effects of the program, two key aspects of evaluating of effectiveness (Datta & Perloff, 1979).

Subsequent reauthorizations of the EHA (P.L. 98-199 and 99-457) were more experimental and discovery-oriented than P.L. 94-142, supporting a diverse array of innovative programs in an extensive search for effective methods of service delivery. One area of programming, the transition initiative, focused on the improvement of secondary curricula and the successful adjustment of students as they move from school to adult life. Since 1984, over 200 model demonstration projects in the area of transition have been funded by the Office of Special Education and Rehabilitation Services (OSERS).

The legislation set similar broad goals for the accompanying evaluation activities mandated by the Act. In addition to mandating the monitoring of accountability factors such as progress in implementation of the Act, program management, and program administration, the law strongly advocated in favor of evaluation of the impact and effectiveness of local and state model demonstration projects in transition with the goal of findings "what works" in the delivery of transition services (Kaufman, 1985). This advocacy was apparent in the mandate that all model demonstration projects funded under the Act collect and report evaluation data concerning the projects' effectiveness. When aggregated

at the national level, these data were intended to give an accurate picture of transition services in the United States, to inform policy and funding decisions as they arose in Congress, and to help local service providers develop and improve their own programs. In 1985, realizing that the demonstration projects might require assistance to obtain the type and quality of data necessary for this initiative, OSERS funded the Transition Institute at the University of Illinois to provide evaluation technical assistance to funded projects and to conduct a program of research on evaluation. This paper is based on data collected by Institute evaluation technical assistance staff during the last five years.

Purpose of the Paper

This paper examines the issues that arose when a particular set of federally-funded model demonstration projects, OSERS' secondary/transition projects, were obligated by federal initiative to improve the amount, quality, and use of evaluation data. These issues are organized into two areas: 1) the nature of federal expectations and needs for evaluation data, and 2) the extent of local capabilities to meet those expectations and provide evaluative data. Regulatory and methodological implications of the discrepancy between federal expectations and local capabilities along with implications for evaluation technical assistance are discussed.

Method

Request for proposals, evaluation plans, evaluation technical assistance needs assessments, technical assistance contracts with individual projects, and final evaluation reports from over 150 model demonstration projects in transition (representing all projects funded from 1987-1990) served as the major data sources in this study. These documents were reviewed and coded by a three-member panel to determine: 1) the types of evaluations requested by the funders and proposed by the projects, 2) the nature of evaluation data to be collected, and 3) the problems that the projects were encountering when trying to implement these plans.

Secondary data sources included surveys from a sample of 109 project directors. The surveys collected information on project directors' perceptions of: 1) federal evaluation purposes, local evaluation purposes, and the degree of concordance between the two; 2) the capability of local programs to produce evaluation data to meet both purposes; 3) the extent to which evaluation data was used by local projects; and 4) factors that enhance or diminish the use of evaluation data by local projects. Surveys were sent to project directors beginning their third and final year of funding (109 out of 157 returned; 70% return rate). Descriptive tables were used to analyze survey data.

Results and Discussion

The Nature of Federal Expectations and Needs for Evaluation Data

As stated in the legislation and accompanying regulations, the federal government needed evaluation data from the model programs for many purposes. The data would be used to give a picture of the current state of transition services in the United States. They would identify areas of success and failure, excess and need. Given the proposed budget cuts and reauthorization of the two major pieces of legislation affecting youth with handicaps—the Education of the Handicapped Act (EHA) and the Rehabilitation Act (RA), OSERS was under extreme pressure to produce data to inform policy and funding decisions as they arose in Congress. Finally, information on successful transition projects would be disseminated to state and local agencies and replication sites would be started, thereby improving service delivery across the country. These needs are evident as we examine the following reviewers' criteria for rating the evaluation section of each application for funding, excerpted from OSERS' requests for proposals for the secondary/transition funding competitions:

> 6. EVALUATION PLAN (0-10 points)
> Quality of the evaluation plan for the project. Look for information that shows methods of evaluation that are appropriate for the project and, to the extent possible, are objective and produce data that are quantifiable. Also, look for information that shows the quality of the applicant's plan to evaluate annually: (1) the grantee's progress in achieving the objectives in its approved application,

(2) the effectiveness of the project in meeting the purposes of the program, (3) the effect of the program on persons being served by the project . . . .

D. EVALUATION PLAN (maximum 10 points)
Review the application for information that shows --
(1) The method of evaluation that is appropriate, objective, and produces quantifiable data.
(2) The significant observable, measurable results expected to be achieved during the project period.

EVALUATION PLAN
The evaluation plan should address the following points:

1. Who is to do the evaluation;
2. What data will be collected, how will they be collected, verified, and stored for ready accessibility;
3. How will data be analyzed and how will the analysis show progress toward and achievement of objectives;
4. State all evaluation criteria which will be used;
5. To the extent possible, the evaluation methodology is expected to include such elements as client and employee reaction to the program; the extent to which skills have been learned by the clients; and the extent to which job performance has been improved or otherwise changed as a result of the training; or the extent to which clients have been returned to gainful employment or assisted to live independently in the community.

If these excerpts are representative of the federal government's expectations regarding the types of evaluation data that were to be gathered at the local level by the secondary/transition projects, federal interest was concentrated upon outcome measures, quantifiable if possible, that indicated the degree to which a program had achieved its stated goals. Results from the analysis of final reports are shown in Table 1. In the case of the model demonstration programs in transition, stated goals of service demonstration were likely to be expressed in terms of outcome variables at the individual level such as: number of students served, student exit status, or student progress in terms of skills gained. Increasingly (14% - 29%), projects included measures of consumer satisfaction with the program. Despite the clear emphasis on outcomes in the federal evaluation criteria, 74% or more of the projects collected information on program characteristics such as instructional strategies used; planning processes; and mechanisms for coordination and interagency

service. About one-fourth (20% - 28%) of the projects collected data on the extent to which these characteristics were actually implemented or operational in the program and an even smaller number (14% - 16%) collected data on the extent and effectiveness of replication attempts.

From this analysis of the types of data collected by the model demonstration projects, it appears that they were strongly influenced by the RFP language, focusing their data collection on traditional quantifiable outcome variables. This trend is further supported by an examination of the types of evaluation approaches or designs used by the projects presented in Table 2. The majority of the projects (53% - 56%) chose a summative, goal-based approach, intent on measuring the extent to which objectives were achieved. About one-fourth of the projects (22% - 31%) took a more formative, decision-making approach to evaluation, using evaluation data to influence program management and development. Relatively low percentages of the projects chose professional review or case study evaluation designs.

The Problem with Emphasizing Outcomes

Given the experimental nature of the model demonstration projects funded under P.L. 98-199 and P.L. 99-457, the restriction of evaluation design to an emphasis mainly on outcome variables runs counter to local circumstances. Outcome data may be most appropriate to describe the current state of transition services on a national basis, but they tell little about the characteristics of the projects themselves. When a major goal of the evaluation initiative was to identify and replicate "successful" transition projects, thereby improving service delivery across the country, the absence of federal interest in implementation and process data seemed contrary to that mission.

Table 1. Type of Evaluation Data Collected by Federally-Funded Model Demonstration Programs in Transition

	Percent of Projects Collecting Data				
	1987 N=36	1988 N=38	1989 N=41	1990 N-41	TOTAL N=157
Number and type of students served	86%	71%	73%	76%	76%
Student progress (skills gained)	61%	61%	68%	68%	64%
Student exit status	47%	50%	61%	61%	55%
Consumer satisfaction	14%	16%	22%	29%	20%
Program characteristics	75%	74%	83%	83%	78%
Program implementation level	28%	21%	24%	20%	23%
Program replication	14%	16%	14%	14%	14%

Table 2. Evaluation Approaches Used by Federally-Funded Model Demonstration Programs in Transition

	Percentage of Projects Using Approach				
Evaluation Approach	1987 N=36	1988 N=38	1989 N=41	1990 N=41	TOTAL N=157
Goal-based	53%	53%	56%	56%	55%
Decision-making	28%	31%	22%	24%	26%
Professional review	11%	8%	12%	12%	11%
Case study	8%	8%	10%	8%	8%

Formative vs. summative. It is natural to assume that newly developing programs such as these model demonstrations, evaluation information should be readily available and targeted toward program implementation and improvement. This implies a formative, flexible design oriented toward program characteristics and process. The federal emphasis on summative outcome-based designs runs counter to the information needs of new projects.

Overpromising and changing goals. A second problem with the requirements of preordinate evaluation designs to measure the attainment of program goals is the tendency of projects to over-promise both activities and outcomes when they submit a grant proposal. This inflation at the start of a project immediately invalidates the use of stated goals as criteria for program effectiveness.

Goal inflation and the innovative nature of the projects (most often proposals are written well in advance of program implementation, with little understanding of the demands and barriers of program) often results in a considerable amount of change in the projects from proposal to final implementation. Table 3 represents our efforts to document the nature and extent of that change.

Referring to the first section of Table 3, when project directors in their final year of funding were asked to compare project accomplishments to date with proposal promises, only 19% (option c) felt that stated goals accurately depicted program accomplishments. Over 50% (options b and d) felt that some stated goals were unlikely to be met during the course of funding. Sixty-three percent (options a and b) felt that they had accomplished things that had not been described in the initial proposal. Therefore, for 81% of all model demonstration projects (all but option c), an evaluation design based upon objectives represented in the proposal may not be useful in assessing actual accomplishments of the project.

In trying to assess the extent to which changes occur and the timing of these changes, we asked project directors to estimate how much their project had changed since the day the

Table 3. Project Directors' Reports on the Extent of Change in Project Goals and Accomplishments (N=109)

	% of Respondents
If you were to take a full inventory of your project accomplishments to data and held it up to proposal promises, how would they match?	
a. Accomplishments diverse, some unexpected, and too numerous to count; goals fully being met.	29%
b. Accomplishments diverse, some unexpected, and too numerous to count; some stated goals unlikely to be met.	34%
c. Accomplishments nicely covered by stated goals which are fully being met.	19%
d. Accomplishments nicely covered by stated goals, some of which are unlikely to be met.	18%

Here's another difficult one. We want to know how much your project has changed from the way it was described in the proposal. Think of it just as it is now. (The percentages represent the median values of respondent estimates.)

a. From a month ago it probably is 100% the same.

b. From one year after funding it is about 95% the same.

c. From six months after funding it is about 85% the same.

d. From the first day of OSERS funding it is about 75% the same.

e. From the day the proposal was submitted, the idea is about 75% the same.

Give us a rough guess as to the percent of the total change that would reasonably be thought of . . . (The percentages represent the median values of respondent estimates.)

a. as conceptual advancement	30%
b. as refinement in detail	30%
c. as responses to trouble	25%
d. as something else	15%
TOTAL	100%

proposal was submitted. The responses to this question are reported in the second section of Table 3. It appears that most project change occurs (20%) during the first year of funding and that the median rating for change was 25%, that is, the project is about 75% the same as described in the proposal.

Finally, project directors were asked to describe the nature of that change in percentages distributed across four categories: conceptual advancement, refinement in detail, responses to trouble, or something else. Thirty percent of the changes were described as conceptual advancement, that is, gaining greater insight or additional substantive knowledge that affected program design. For example, one project completely changed its strategy for involving parents as a result of a needs assessment that caused them to reconceptualize the roles of parents from one of caretaker to one of advocate. An additional 30% of change was described as refinement in detail. In this case, the stated goals of the program may not be greatly affected; just broken down and stated with greater specificity. Twenty-five percent of change was interpreted as responses to trouble such as social and economic factors; lack of public participation, poor timing of funding; failure of interagency cooperation; shortage of trained personnel and administrative problems. The remaining 15% of change was ascribed to other factors such as personnel turnover, changes in needs of target population, changes in organizational climate and so on.

The significant amount of change that occurs from proposal to implementation and the fact that some stated goals are not achieved by projects while other unanticipated accomplishments occur mitigate against the use of preordinate goal based designs.

Funding cycles. A final problem associated with the federal focus on outcomes has to do with funding cycles. The three year funding cycles that dictate the timetable for project evaluation may not be long enough to allow for attainment of stated goals, never permitting projects to understand fully the impact of their intervention.

The need for formative, process-oriented information by newly developing projects, the invalidity of stated goals as a basis for evaluating program effectiveness, and limitations of a

three year funding cycle mitigate against the use of summative, outcome oriented evaluations emphasized by the federal government.

The Problem with Emphasizing Quantitative Data

The emphasis on the collection of quantitative data compared to the focus on outcomes was easier to understand in theory, though not in practice. If the intent was to aggregate and compare locally produced findings, it is presumably easier to do so with quantitative data than with narrative or qualitative reports. Unfortunately, there are also substantial problems associated with the aggregation of quantitative data. Those who know the pitfalls of trying to compare nonequivalent groups are familiar with the limitations of this plan.

The absence of specific guidelines in the RFP to standardize the types of evaluation data to be collected permits creativity in evaluation design on the part of the projects, but exacerbates problems associated with cross-site aggregation of evaluation data. Use of different instruments across projects to measure student and program characteristics make combination or comparison of data problematic. Under-operationalized or differing definitions of disability groups, exit status, and skill acquisition make interpretations of aggregated data difficult. If the federal government has the responsibility to evaluate the secondary/transition projects as a whole, then a system for standardizing the collection of certain evaluation data may be desirable to enable the meaningful aggregation of data across sites. This standardized data may suffice as the only requirement to fulfill the government's evaluation expectation, but given the importance of understanding the individual workings of each project, a more desirable alternative may be to request the collection of a small subset of standardized data as an adjunct to the projects' individual evaluation design.

There are other concerns with the use of a quantitative approach to measuring change in special populations. The quantitative emphasis reflects a traditional approach to educational evaluation, namely, the use of quantitative student progress information (i.e., achievement test scores) as an indicator of program effectiveness. The narrowness of this

approach has been problematic in the evaluation of educational programs for decades (Stake, 1986). When applied to the secondary/transition projects, the utility of this approach is reduced even further by the difficulty of determining levels of "adequate" progress for an exceptional population as well as the lack of reliable and valid instrumentation to measure the outcomes sought by the projects. The long-term outcomes of transition such as community employment, increased independence, and greater lifestyle satisfaction are not easily assessed. In some cases instruments have been developed to quantify some of these outcomes (Heal & Chadsey-Rusch, 1985; Keith, Schalock, & Hoffman, 1986; Landesman, 1986). In most cases, however, widely accepted assessment instruments are not available and, in fact, for some aspects of the complexities of transition services, outcome indicators may not be known.

We have discussed the problems associated with outcome-oriented goal-based assessment, and quantitative measures, however, the nature of federal expectations and needs for evaluation data represents only one side of the dilemma. On the other side we have the model demonstration project charged with the responsibility of providing evaluation data along with the more immediate responsibility of providing services at a local level. Their perceptions of evaluation and their ability to respond to the federal evaluation initiative are discussed in the next section.

The Extent of Local Capabilities to Provide Evaluation Data

To understand the extent to which local project personnel are capable and willing to meet federal expectations it is necessary to consider: 1) evaluation expertise at the local level; 2) local need for evaluation information; and 3) the rewards associated with responding to federal and local needs.

Evaluation and Expertise at the Local Level

This variable becomes important with the discovery that, due to limited funds set aside for evaluation, in the majority of the projects, the task of evaluation fell to the project director and the program staff (Table 4).

Table 4. Personnel Primarily Involved in Evaluation in Federally-Funded Transition Model Demonstration Programs

	1987 N=36	1988 N=38	1989 N=41	1990 N=41	TOTAL
Project staff	87%	86%	84%	79%	
Advisory Board	10%	11%	9%	12%	
Third party evaluator	3%	3%	7%	9%	

Objectivity. Placing the responsibility for program evaluation under the jurisdiction of the project director raises several issues associated with the validity and objectivity of any evaluation planned and conducted in this matter. Of primary concern is the ability of the project directors to evaluate objectively their own programs. Project directors believe deeply in the validity and worth of their program. They are vitally involved in seeing that their programs succeed. In many cases, project directors derive security from an identification with the organization and looks to it as a source of personal advancement and recognition (Gurel, 1975). Although project directors may have the most complete understanding of the goals and activities of their projects, the extent of their personal involvement may limit their ability to collect and communicate objective evaluation data. Limitations may be minimized when the evaluation is conducted for internal use in staff development or program improvement; they may be exacerbated in the case of reporting the extent of program success to an external funding agency such as the federal government.

Expertise. In addition to the problem of objectivity, some project directors reported that they did not possess the skills and training necessary to conduct a program evaluation. Like the projects they administer, the project directors were a diverse group, representing a variety of disciplines, educational levels, and professional experiences. In an article on evaluating employment services, Schalock and Hill (1986) observe:

> If program managers were asked whether they were trained to do the tasks they are currently doing, the majority would most likely answer "no." They would add that they were not trained in program monitoring and evaluation, the terms were not in their lexicon five years ago, and that program monitoring and evaluation are still not their favorite activity. If one adds the lack of comprehensive data systems and the confusion over operationalizing and quantifying factors that some consider to be "non-quantifiable," it is no wonder why program personnel become frustrated when asked to provide outcome data of their programs' effectiveness, efficiency, and benefit (Boschen, 1984).

Local Need for Evaluation Data

By placing the responsibility for the design, conduct, and reporting of evaluation activities at the local level, problems of accuracy and reporting, varying levels of evaluation expertise, and lack of systematicity are all but unavoidable. However, local control creates the opportunity for project directors and their staff (with the necessary assistance) to obtain evaluative information that is most meaningful to them; that which can lead directly to program improvement and informed decision making. Unfortunately, this information is usually gained through study of the implementation and processes of a program and the effectiveness of particular program components and not by assessing overall effectiveness as advocated in the federal guidelines (Leviton & Hughes, 1981; Smith, 1991).

The Rewards Associated with Responding to Federal and Local Needs

Given the limited amount of personnel and monetary resources allocated for evaluation, it may be unlikely that a single locally conducted evaluation can meet both local and

federal needs to the fullest extent. Under these circumstances, the project director may be put in the position to give priority to either local or federal needs or to address both superficially, leaving one or both users of evaluation information short changed.

From the point of view of the project directors, the argument in favor of choosing the local perspective was a strong one. The rewards of effectively responding to local evaluation purposes often had a direct and profound impact on the project (and its director). As a direct result of an evaluation, administrators at an overseeing agency may have been made aware of the merits of one of its programs, a local funding source may have agreed to pick up funding at the end of an OSERS grant period, or a political or advocacy group may have lobbyed for legislation or funding to support a demonstrably effective program.

The rewards of providing evaluation data at the federal level were less clear. Although the federal government's right to these evaluation data was never questioned, the fact that the federal government was not seen as a primary evaluation audience by most of the projects indicates that the accountability link between federal funding and program evaluation was not firmly established in the projects' minds. On the average, the evaluation section of a grant application is worth five to ten points out of 100 in the review process. In most cases, scores of four or five were awarded to even the most meager evaluation plans. Evaluation data were to be included in each continuation proposal, but continued funding was typically granted in the absence of these data. Regulations required that the final report issued at the end of program funding include the results of the completed evaluation study, but no funding was contingent upon the provision of that data and final reports often did not contain evaluation data.

Conclusion

There is a lot of work to be done to narrow the gap between the two consumers of evaluation discussed in this paper, the federal government and local project directors. Needs for different types of evaluative information, unclear rewards for complying with

federal evaluation initiatives, questionable feasibility and utility of an emphasis on outcome evaluation, inflated program objectives, and incompatible reporting cycles are all problems that contribute to the difficulties encountered in federally mandated reporting of local evaluation data.

Change is needed in at least two areas: the provision of evaluation technical assistance and the use of evaluation methodology by the model demonstration programs.

Implications for Evaluation Technical Assistance

Early intervention. The lack of evaluation expertise among project directors, the limited amount of funding available for evaluation at the local level, and the complexity of the evaluations to be carried out come together to create a situation of great need for evaluation technical assistance. Currently, technical assistance begins at the moment of funding. It may be more effective to take a preventative approach, offering evaluation technical assistance to persons who are in the process of preparing their proposals. Delivered in this manner, evaluation technical assistance could focus upon stating realistic goals, building flexibility into evaluation design to accommodate change, and addressing federal information needs.

Evaluation needs assessments. Immediately upon funding, evaluation needs assessments could be conducted with all new projects, providing the first opportunity for readjustment of the evaluation plan. These needs assessments can serve as the basis of ongoing technical assistance activity.

Continued involvement with projects. Continuing inservice on evaluation should be offered for project staff who are responsible for evaluation. This contact might involve the review of actual evaluation materials and data from the projects, instruction in evaluation methods, and guidelines for developing the final report. Though labor intensive, provision of evaluation technical assistance increases capacity among project staff to evaluate future projects and promotes the use of evaluation among model programs in general.

Implications for the Regulation of Evaluation by the Federal Funders

Analysis of the evaluation sections of several RFPs revealed the federal emphasis on preordinate, goal-based, quantitative evaluation designs. Argument has been made that for many reasons, evaluations of this type may not be best suited for model demonstrations in transition. Nevertheless, the presence of the language in the RFP strongly influences the type of evaluation conducted by the projects. Reform at the federal level should involve three elements: general recognition that evaluation cannot serve all goals; use of the continuation proposal to realign goals and build flexibility into evaluation designs; and allocation of a percentage of project budgets for evaluation among model programs.

General recognition that evaluation cannot serve all goals. The intention of aggregating and using locally produced evaluation data to represent national trends is likely to be impossible without some standardization of design, instrumentation, and specification of variables of interest. Perhaps a compromise between local and federal needs can be met with a two-tier evaluation design. At tier one, a small set of outcome information, specified and operationalized by the federal government, could be collected and summarized with the goal of producing summative information. At tier two, local projects would be encouraged to plan an evaluation to meet local needs and to produce information to aid replication. Tier two evaluation would likely emphasize formative, process-oriented measures and rich description of context.

Use of the continuation proposal to realign goals and build flexibility into evaluation designs. Projects are required to submit continuation proposals after each year of funding. At present, these proposals emphasize fiscal rather than substantive changes and contain no language related to evaluation. It seems that by encouraging projects to account for project change, realign goals, and report upon evaluation findings or changes in the evaluation plan in the continuation proposal, a mechanism exists to account for and promote the accommodation to change in project goals.

Allocation of a percentage of project budgets for evaluation among model programs. Without sufficient funds allocated for evaluation, there is little hope of improving its quality at the local level. If projects were asked to earmark 5-10% of their operating budget for evaluation, the use of external evaluators or evaluation consultants and the amount of time allocated to evaluation at the local level would be likely to increase proportionately.

The findings of this study have added to understanding the context of evaluation at the local level in relation to federal expectations. Continued work is necessary to assess the extent to which technical assistance can be useful in resolving the conflict between local capabilities and federal expectations along with the development of strategies for negotiating, providing, and evaluating technical assistance efforts.

References

Boschen, K. A. (1984). Issues in evaluating vocational rehabilitation programs. Rehabilitation Psychology, 21(1), 37-48.

Datta, L., & Perloff, R. (1979). Improving evaluations. Beverly Hills, CA: Sage.

Gurel, L. (1975). The human side of evaluating human service programs: Problems and prospects. In M. Gettentag & E. Struening (Eds.), Handbook of evaluation research. Beverly Hills, CA: Sage.

Heal, L. W., & Chadsey-Rusch, J. (1985). The Lifestyle Satisfaction Scale (LSS): Assessing individuals' satisfaction with residence, community setting, and associated services. Applied Research in Mental Retardation, 6, 475-490.

Kaufman, M. (1985). Closing comments. Proceedings of the 1st Annual Project Directors' Meeting, Washington, DC.

Keith, K. D., Schalock, R. L., & Hoffman, K. (1986). Quality of life: Measurement and programmatic implications. Lincoln, NE: Region V Mental Retardation Services.

Landesman, S. (1986). Quality of life and personal satisfaction: Definition and measurement issues. Mental Retardation, 24(3), 141-143.

Leviton, L. C., & Hughes, E. F. X. (1981). Research on the utilization of evaluation: A review and synthesis. Evaluation Review, 5, 525-548.

Schalock, R. L., & Hill, M. L. (1986). Evaluating employment services. In W. E. Kiernan & J. S. Stark (Eds.), Pathways to employment implications for the future (pp. 285-302). Baltimore: Paul H. Brookes.

Smith, N. (1991). Needed research on optimizing the structure of evaluation designs: The issue of flexibility in evaluation work. Advances in Program Evaluation, 1A, 56-68.

Staff. (1986, February 19). House seeks extension of EHA, while Senate wants full reauthorization. Education of the Handicapped, 3.

Staff. (1986, February 19). Reagan's 1987 budget would set P.L. 94-142 grants at 1985 levels. Education of the Handicapped, 1.

Stake, R. E. (1986). Quieting reform. Urbana: University of Illinois Press.